武汉大学面向本科生的数据素养通识课程的建设与思考*

■ 刘霞 方小利 郑怡萍
武汉大学图书馆 武汉 430072

摘 要:[目的/意义]旨在通过分享武汉大学数据素养通识课程的建设过程,为同行开展相关工作提供参考。[方法/过程]通过案例分析、数据统计和数据挖掘可视化方法对"武大通识3.0"背景下的《数据素养与数据利用》课程的教学目标、教学内容、教学特色以及教学效果进行分析,探索课程体系的合理性。[结果/结论]授课方式和课内外实践内容的设置能够很好地调动学生学习的积极性,提高学习效果;课程考核结果、学生学习前后的问卷调查及课后反馈表明,课程学习明显提升了学生的数据素养。

关键词:数据素养 通识课程 本科生 课程建设 武汉大学
分类号:G254.9
DOI:10.13266/j.issn.0252-3116.2020.22.007

大数据时代,数据对日常生活、科学研究和社会发展均具有重大影响。日常生活中,人们是数据的利用者,亦是有意或无意的数据生产者,数据素养成为影响人们生活节奏和品质的重要技能;科学研究中,数据既是科学发现的基石和源泉,也是检验科学研究价值的试金石[1];社会需求方面,数据分析与运用能力成为各个领域发展的驱动力,用人单位需要的不仅仅是数据分析师,还需要具备数据素养的各专业人才[2],数据素养是大学生就业的关键技能[3]。该背景下,加强大学生的数据素养教育,已经成为业界共识。

为深化教育教学改革,武汉大学2017年起推出了"武大通识3.0"计划[4],重构通识课程体系,要求除公共基础课外的所有通识类课程都必须重新申报,通过评审的课程方可纳入培养方案。图书馆申请了一门全新的课程——《数据素养与数据利用》,并获批立项,且被纳入为学校的创新创业课程。本文基于已进行的2019-2020学年两学期课程开展情况,介绍该课程的教学目标、教学内容、教学特色以及教学效果,旨在通过对课程的系统总结,为同行提供参考。

1 国内外研究与实践进展

1.1 国外概况

2006年起,美、英等国国家级基金相继资助了一些数据素养教育研究项目[5],这些项目分别面向中学生、本科生、研究生和科研人员等不同群体展开需求调研,确定核心能力,设计课程体系,开展教学探索,积累了丰富的理论成果和最佳案例[6-10]。此后,数据素养教育迅速普及,大学图书馆成为实施教学的主体,采用学分课程、嵌入式教学、在线教育、培训讲座或专题研讨会等多种形式,面向特定学科[11-12]和/或特定对象[13-14]开展数据素养教育。虽课程大纲和内容体系各有差异,但主要围绕科研数据生命周期展开,具体涉及下述多个或全部知识和技能:数据管理基础知识、数据管理政策和计划、数据类型和数据格式、数据命名与文件组织、数据描述与元数据、数据质量控制、数据保存、安全与备份、数据遴选与评估、数据分析处理技术及工具、数据共享与长期保存、数据发现与检索、数据引用、数据伦理与法律问题[15]。

国外专门针对本科生的课程不多,且多采取嵌入式教学方式,以提升本科生的数据利用能力为目的。如普渡大学图书馆嵌入到理工科(STEM)本科生的两门专业必修课中去,以讲授部分课程或负责相关课程作业的形式,培养学生发现、评价、获取数据的能力[16];加州大学洛杉矶分校将数据素养内容嵌入到社会学信息素养课程之中,通过两次课来培养学生阅读、评价及正确引用数据表格,运用工具创建表格,以及结

* 本文系"武大通识3.0"项目(数据素养与数据利用)(2019年9月-2022年6月)(项目编号:武大本函[2018]158号)研究成果之一。
作者简介:刘霞(ORCID:0000-0002-0624-4394),研究馆员,博士,E-mail:liuxia@lib.whu.edu.cn;方小利(ORCID:0000-0001-6765-5288),馆员,博士;郑怡萍(ORCID:0000-0002-2300-7729),馆员,硕士。
收稿日期:2020-06-07 修回日期:2020-09-04 本文起止页码:69-76 本文责任编辑:易飞

合上下文理解图表的能力[17]。

1.2 国内进展

国内相关研究可以分为 3 类：

（1）针对美国、英国、加拿大等国的数据素养教育实践，从教学对象、教学目标、教学形式、教学内容、教学评估等角度展开全面调研[15,18-19]，或典型案例分析，总结其理论基础和实践经验[20-21]。

（2）对数据素养教育进行理论探讨和策略研究，涉及目标定位与体系构建[22-23]、教育模式[24]、内容框架与课程体系[25-27]、开展策略[28-29]等。

（3）对国内数据素养教育实践展开调查。张新红[30]通过对国内外 14 所高校的调研和比较，认为国外已经构成完整的数据素养教育体系，而国内则大多依托于信息素养教育。宋嘉丽等[31-32]调查了 42 所“双一流”建设高校，发现各校虽然开展了多种形式的教学培训，也涉及了科学数据基本知识、数据库使用、数据处理、数据分析和数据应用等数据素养相关内容，但未从数据生命周期整体流程入手，也未有高校开展科学数据素养专题课程。

综上所述，数据素养教育在国外已经非常普及，并形成了面向不同层次不同学科的多样化的教学培养体系；国内则更多地停留在理论探讨阶段，具体实践案例少见报道，因此，本文希望及时总结《数据素养与数据利用》课程的建设情况，为我国高校图书馆开展数据素养通识教育提供参考案例。

2 课程目标与内容体系建设

2.1 教学目标

教学目标和内容的设定应该充分考虑选课对象的实际需求，同时应具备一定的前瞻性。

为了解需求，本课程在课前进行了问卷实名调查。问卷从数据意识、数据知识、数据能力、数据伦理道德等角度设计了 20 余个问题。调查发现，绝大多数本科生有过数据利用经历，其应用场景涉及课程学习、学生工作、日常生活、科研项目、论文写作、活动比赛、社会实践等各个方面（见图 1），其主要关注点是如何获取和利用数据，因此有必要将教学重点定位于数据的发现、采集、处理和分析能力的培养；与此同时，大学生亦是部分数据的生产者，尤其是研究型大学的本科生，毕业时多数选择升学或出国深造[33]，未来将不同程度地走进科学研究，因此在提升个体基础数据素养的同时，有必要加强基于科研的数据素养教育[34]。事实上，开课问卷调查也表明，同学们普遍缺乏数据应用技能和相关的数据知识，如利用工具进行异构数据集的处理、了解本学科通用的数据集描述方法、判断并清洗数据、数据及其生命周期相关知识、数据分析工具的掌握、数据背景信息的同步记录等。

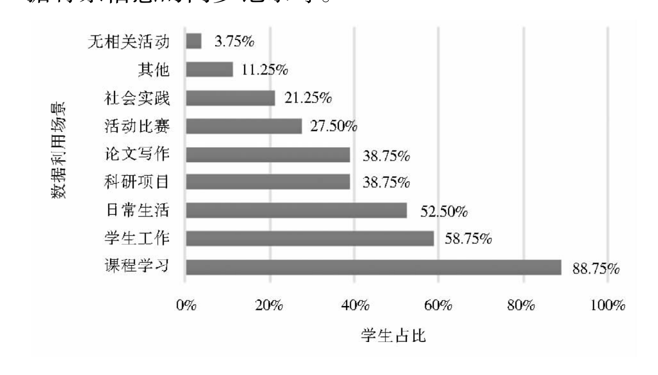


图 1 选课学生数据利用场景分布

基于上述分析，本课程的教学目标定位为帮助大学生自发建立数据价值意识，掌握获取数据、分析评价数据、管理和利用数据的途径和方法，培养其在学习和实践活动中对于数据问题的应对能力，亦为科研创新奠定基础。

2.2 教学内容及计划

根据教学目标，本课程主要从数据利用者的角度，按照数据生命周期，分成 4 个阶段设置教学内容：

（1）数据获取：涉及一手或/和二手数据，包括馆藏数据、政府数据、科研数据及其他互联网数据四类数据资源；

（2）数据处理：包括数据评估和数据清洗 2 个阶段；

（3）数据分析：包括方法选择和结果呈现 2 个步骤；

（4）数据发布：包括正式出版和非正式发布 2 种方式。

其中，前 3 个阶段可能会根据需要多次重复，直至获得满意的结果，才进入数据发布阶段。此外，整个数据生命周期，都需要进行过程管理，具体包括数据文件命名、格式选择、存储与备份，数据背景信息的实时记录或元数据描述，以及数据的全周期质量控制等，见图 2。

基于这个生命周期，参考美国数据信息素养项目（DIL）开发的核心数据素养能力及其相关数据技能[27]，本课程制定了为期 11 周 32 个学时的教学计划，如表 1 所示。其中，数据基础知识部分讲解占 5 个学时，学生的学习交流研讨占 3 个学时，这两部分共分 3 次课进行，其余课程内容则按照 4 个阶段展开教学，分 8 次授课。

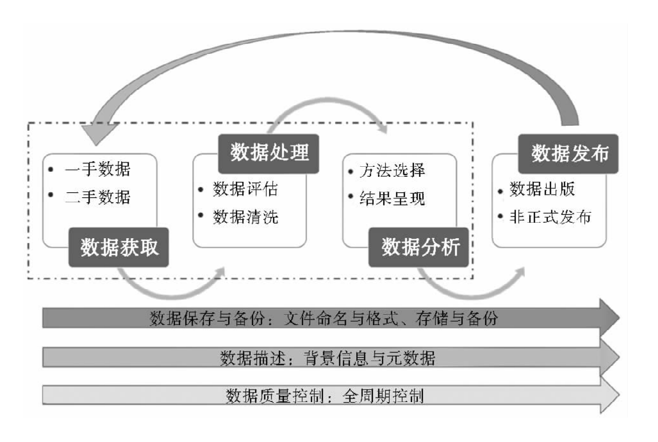


图 2　数据生命周期——利用者角度

3　教学实施与特色

本课程 2019 年 9 月首次开课，至今已开课两学期，选课学生总计 96 人，其年级和学科分布如图 3 所示。97% 为 1－3 年级同学，其中大二占比最高（37%）；学科分布以社会科学居多（45%），工学、信息科学、理学和人文科学占比接近，医学部因校区较远选课学生少。总体而言，选课学生的年级和学科分布合理，体现了通识课的特点。

表 1　《数据素养与数据利用》教学计划

内容安排	授课知识点	课堂练习/讨论考察的知识点	课后阅读/学习资料
数据基础知识 第 1－2 周 （5 学时）	• 开课背景及课程安排 • 数据与科学数据 • 数据管理概述 • 数据过程管理要点	• 开课调查：数据素养调查 • 课堂练习（5 个）：数据概念查找、数据识别与归类、数据文件格式、数据文件命名、数据生命周期 • 课堂讨论：数据描述的重要性	• 参考书目：7 种 • 论文：5 篇 • 文件/标准：2 个 • 网站/网页：4 个
数据获取 第 3－5 周 （9 学时）	• 数据资源介绍：政府数据、科研数据、馆藏数据 • 数据采集：互联网数据采集、主题数据采集、采集工具介绍	• 课堂练习（7 个）：各类平台的数据查询、各类主题的数据检索 • 课堂讨论：数据采集经验分享	• 网站/网页：16 个 • 报告：1 个 • 补充学习 ppt/视频：5 个
期中报告	课后作业：参照模版，整理形成学科相关数据资源目录（包括平台名称、平台介绍、涉及本学科的资源、获取方式等）		
数据处理 第 6－7 周 （6 学时）	• 数据质量评估 • 数据完整性校验 • 数据清洗：概述及方法 • 清洗工具介绍	• 课堂练习（7 个）：质量评价维度和方法、脏数据分析、数据完整性校验、数据清洗流程、数据内容提取	• 网站/网页：1 个 • 论文：4 篇 • 补充学习视频：2 个
数据分析 第 8－9 周 （6 学时）	• 数据分析概述 • 数据分析工具 • 数据可视化	• 课堂练习（3 个）：数据分析分类、统计分析工具使用、数据挖掘	• 工具使用视频：5 个 • 电子书：2 本
数据发布 第 10 周 （3 学时）	• 数据引用 • 数据发布 • 学术规范 • 隐私保护	• 课堂练习（1 个）：寻找引用实例 • 课堂讨论（3 个）：认识数据论文、数据伦理、个人隐私保护经验分享	• 文件/规范：6 个 • 论文/报告：3 篇 • 网站/网页：1 个
期末交流 第 11 周 （3 学时）	• 组成团队（5－6 人），结合专业或兴趣选题，进行数据采集获取、数据处理、数据分析，最后形成数据研究报告（Word/PDF 及 PPT） • 各小组代表上台展示：选题思路、数据源及采集方法、清洗和分析过程、结论等 • 学生相互评分或点评，教师点评 • 完成结课问卷调查：本科生数据素养调查		

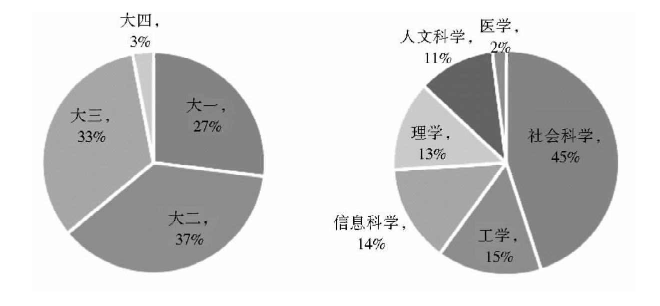


图 3　选课学生年级及学科分布

本课程第一期为实体课堂教学，第二期因新冠疫情改为录制视频线上教学，两期的教学方式虽然不同，但教学理念和教学内容相同。经过两期的教学实践，课程具有以下特色：

3.1　引入多种参考资料，教学资源丰富

因缺少专门教材，为保证学习效果，一方面挑选了 7 本有馆藏的参考书供学生选读，内容涉及基础数据素养、科研数据管理以及统计学、数据分析、数据科学等领域入门级教材；另一方面，如表 1 所示，每章均提供补充学习和阅读资料，包括课件、政策文件、标准规范、相关网站、学术论文和专题报告等，部分重要内容列为必读清单。此外，结合授课内容，还制作了 8 个实操类微视频供学生课后自学，如数据清洗演示、SPSS 操作入门、用 Tableau 创建词云图等。

3.2　课堂教学理论与实践相结合

从时间分配来看，表 1 的授课知识点为授课内容，教学视频总计 802 分钟，约 18 学时（占比 56%），在系

统介绍相关知识的基础上,也结合案例讲解典型的方法和工具,为未来的自主学习打下基础;另外,44% 的课时主要用于课堂阅读、课堂练习和讨论交流(详见表 1 的“课堂练习/讨论考察的知识点”)。其中课堂练习以巩固所学、提升实操能力为目的。分为测试题和实践题两类,测试题主要检查对知识点的掌握程度,采用选择题和简答题形式;实践题为引导学生上机操作,完成特定任务;课堂讨论则更多地是激发同学们的兴趣,通过个人经历的分享,达到发现问题、分享经验的目的。

3.3 综合运用多种教学辅助工具

第一学期采用实体课堂教学,利用学校网络教学平台珞珈在线(电脑端)/学习通(手机端)分享教学资源、组织课堂互动(签到、测试和作业提交、讨论、问卷调查等),同时组建课程 QQ 群用于信息发布和师生沟通;第二学期采用线上教学,录制教学视频,利用教学平台上传视频供学生课前自学和课后复习,并每周定时 QQ 群直播,集中讲解上节作业或本节重点。根据两学期的教学历程,可以发现利用教学平台辅助实体课堂教学后只需再录制教学视频即可转型开展 MOOC 教学;此外,教学平台可以真实记录学生学习情况,便于教师优化教学。例如线上教学期间仅教学视频总访问量即达到 10 655 次,每个视频都有学生反复收看,最高重复了 6 次,教师可以据此了解学生的兴趣点和难点。

3.4 注重学习过程管理和综合应用能力的培养

本课程的成绩评定采取综合考核方式,考勤占 5% - 10%,课堂互动和作业占 25% - 30%,期中学科数据资源报告占 10% - 20%,期末数据研究报告占 50%,前面三项根据上课形式进行微调。其中,考勤、课堂互动和作业主要考察学生的参与度,是对学习态度的衡量,作业完成的好坏仅供教师教学参考;期中提交的学科数据资源报告可引导学生利用数据发现与采集方法搜集专业相关资源,为其未来的专业学习打下资源基础;期末的数据研究报告及其交流互评则是对同学们的数据综合利用能力、团队协作能力和展示演讲能力的综合考察。此外,课堂引入“共享杯”大学生科技资源共享服务创新大赛[35]和“慧源共享”高校开放数据创新研究大赛[36],鼓励同学们结合大赛选题或者在课程学习的基础上组队参加,在实践中提升对数据的理解和应用能力。

4 教学效果

4.1 学习成果分析

为充分了解学生对知识点的实际掌握情况,我们选取下学期期末学生分组完成的数据研究报告进行分析(如表 2 所示)。在数据源的选择方面,平均每组使用了 3 个不同的数据源,涉及资源类型 1.9 种,主要为馆藏数据库、政府开放数据和互联网资源,表明学生们基本掌握了各种数据源的使用;在数据清洗和分析方面,各组都使用了 Excel,但已经不再将其视为简单的表格制作工具,而是作为一款最常见且功能齐全的数据清理和分析工具来灵活应用。此外,多数小组采用了一种以上的数据清洗/分析工具,包括 OpenRefine、SPSS、Python 等,这也体现了学生们在教师引导下,能够更深入地了解并掌握专业的数据分析工具。

总体而言,同学们能结合专业或兴趣选题,利用各类数据资源采集数据,并选用合适的数据清洗和分析工具,分工合作完成数据研究报告、展示研究成果,且相互之间的点评客观、到位,取得了很好的交流和学习效果。

表 2 2019 - 2020 第二学期数据研究报告

小组学院分布	小组选题(平均得分)	数据源数量/类型数(个)	数据清洗/分析工具
经管学院	基于全球新冠肺炎确诊病例统计数据探究各国疫情应对管理状况及成效(89.8)	2/2	Excel
法学院、历史学院、社会学院、政管学院	经济发展与教育资源的相关分析——以云南、河南、广东三省为例(90.0)	2/1	Excel、OpenRefine
经管学院	世界石油价格及其影响因素分析(97.8)	5/3	Excel
电信学院、遥感学院	近十年我国婚姻状况变化及影响因素分析(93.2)	3/2	Excel、OpenRefine
外语学院、文学院、艺术学院、哲学学院	1978 - 2018 年改革开放以来我国文化产业的发展状况分析(85.7)	3/2	Excel
经管学院	1998 - 2002 年中国工业企业劳动生产率影响因素分析(90.3)	1/1	Stata、Excel
城设学院、信管学院	2009 年 - 2019 年中国本科毕业生就业情况分析(97.3)	3/1	Excel、Python
计算机学院、健康学院、数学学院	1990 年 - 2018 年全国人口医疗卫生状况分析(94.8)	4/2	Excel、SPSS、Matlab
动机学院、弘毅学堂、水利水电学院	2001 - 2018 年中国粮食安全问题研究分析(88.2)	5/3	Excel、1stOpt
土建学院、资环学院	长江水质数据资源研究报告(90)	2/2	Excel

4.2 学生反馈分析

本课程在结课时同样进行了数据素养问卷调查。实体课堂和网络课堂分别有32和48位同学完整参与了课前和课后两次调查,可据此观测教学效果。问卷针对每个问题,提供完全不同意、不同意、一般、同意、完全同意5个选项。我们利用SPSS对4次有效问卷进行了信效度评估,结果表明,4次问卷的平均Cronbach's α系数为0.945,平均KMO值为0.792,并且Bartlett样本球形检验(P<0.001)达标,表明问卷结果可靠。学生选课后各项测度指标的变化情况如表3所示,其中"不同意"包括了"完全不同意"和"不同意"两种情况,"同意"亦然,对应数据为结课时勾选学生所占比例;而"变化"则指相对于开课调查,该比例降低或提升情况。

总体而言,两个班的所有观测指标均得到较大提升,"不同意"从18%降为2%,而"同意"则从45%上升为82%。各项指标上升幅度不一,其中提升最快的指标包括:了解数据及其生命周期相关知识、利用工具进行异构数据集的处理、了解本学科通用的数据集描述方法、根据数据来源判断数据质量、掌握一定的数据分析和可视化工具、了解学科数据来源、会同步记录数据获取/处理相关背景信息,掌握基本的数据统计方法、合理解读数据分析结果、会保存不同阶段的数据文件以便溯源等,提升比例均超过40%,而这些也多数是学生选课之初较为欠缺的知识和技能。可见,课程对学生的数据素养提升效果明显。

表3　开课前后数据素养问卷调查比较

编号	问题/指标	实体课堂		网络课堂	
		不同意(变化)	同意(变化)	不同意(变化)	同意(变化)
1	了解数据及其生命周期相关知识*	0%(-25%)	88%(59%)	4%(-21%)	88%(52%)
2	了解数据具有时效性/经济性/敏感性	0%(-9%)	91%(28%)	0%(-8%)	92%(21%)
3	学习/科研/生活中形成了数据思维	3%(-9%)	81%(41%)	2%(-13%)	81%(21%)
4	能发现和准确表达出数据需求	3%(-9%)	72%(34%)	2%(-13%)	81%(42%)
5	能判断所需数据内部产生或外部获取	3%(-16%)	81%(44%)	2%(-13%)	83%(25%)
6	能确定合适的外部数据获取渠道	0%(-13%)	81%(31%)	4%(-10%)	92%(33%)
7	能根据来源判断数据可靠性/可用性	0%(-13%)	84%(41%)	0%(-8%)	88%(48%)
8	了解本学科的主要数据来源	0%(-16%)	84%(56%)	0%(-17%)	81%(33%)
9	能检验数据的完整性/一致性,并清洗数据*	3%(-16%)	69%(25%)	2%(-29%)	85%(48%)
10	能利用工具进行异构数据集的处理*	3%(-25%)	66%(38%)	4%(-46%)	75%(60%)
11	掌握基本的数据统计方法*	3%(-19%)	66%(31%)	6%(-17%)	71%(46%)
12	掌握至少一种数据分析和可视化工具	3%(-9%)	78%(50%)	4%(-29%)	69%(40%)
13	能对数据合理分析,并形成结论	3%(-13%)	84%(47%)	0%(-8%)	81%(35%)
14	会保存不同阶段的数据文件以便溯源	3%(-9%)	84%(47%)	2%(-21%)	77%(35%)
15	会同步记录数据获取/处理相关背景信息*	3%(-16%)	75%(34%)	0%(-29%)	77%(48%)
16	了解本学科通用的数据集描述方法*	3%(-16%)	69%(50%)	4%(-33%)	75%(50%)
17	有及时存储和备份数据的习惯	3%(-9%)	88%(25%)	0%(-8%)	83%(13%)
18	了解数据采集/利用相关法律和伦理道德	3%(-6%)	88%(41%)	0%(-21%)	94%(38%)
19	能合理获取和使用数据,规范引用	0%(-6%)	97%(13%)	0%(-2%)	90%(4%)
20	愿意共享自有数据	3%(-6%)	84%(19%)	0%(-2%)	88%(17%)
	平均	2%(-13%)	80%(38%)	2%(-17%)	83%(35%)

注:标*的为课前调查中选择"不同意"或"完全不同意"比例超过20%的问题

此外,我们还收集了两学期的学生以小组或个人为单位撰写的结课感想43份,利用Python进行分词分析。图4为屏蔽了"数据""数据素养""数据利用""课程"等与课程名称相关的关键词并设定词频阈值为3后形成的词云图。不难发现,出现频率较高的词汇有"数据分析""分析""知识点""专业知识""数据库""数据处理""清洗""excel""重要性""工具""软件"等,说明学生通过本门课程学习,在数据的意识、数据源、数据处理及分析工具等方面收获较大,这与表3中"能发现和准确表达出数据需求""了解本学科的主要数据来源""能检验数据的完整性/一致性,并清洗数据""掌握至少一种数据分析和可视化工具"等指标的积极改变相吻合。



图 4　学生课程感想词云图

5　总结与思考

《数据素养与数据实践》通识课授课一年，取得了较好的效果，但也存在一些问题需要在后续的教学实践中不断改进和优化。

5.1　教学内容：方法与技能教育的平衡

如前所述，本课程的选课学生覆盖所有年级，涉及几乎所有学部，学科和年级的差异导致学生学习起点的不同，直接影响了学习效果。这主要体现在数据处理和数据分析部分，授课教师着重培养学生的数据思维，因此会系统介绍基本的原理、通用方法和实现工具，并结合案例进行讲解，但是由于课堂教学时间有限，不能对各个工具进行详细深入的讲解。从评教反馈来看，部分学生觉得“听起来比较吃力”“对计算机基础较弱的同学并不友好”，希望“更加通俗易懂”“具体介绍部分软件的学习方法”“课程属于比较硬核的那种”；而部分学生反而认为“课程内容深度应该再加上一点”；这种需求上的冲突需要在实际教学中不断探索可行的解决方案。未来可探索细化教学对象，分年级[37]或专业背景实施部分内容的小班或小组教学。

5.2　教学方式：实体课堂与网络教学的有机结合

新冠疫情使得本课程在短短一年内尝试了实体课堂和网络教学两种方式。从课前课后的问卷调查来看，两种授课方式的学习效果相差无几；从教师和同学们的体验来看，实体课堂在整体进度把控、实操指导和小组交流方面优势明显，而网络教学则在课程内容的丰富度、教学视频的可重复收看、学习过程的可观测性等方面更有利于学生实现个性化学习。因此，未来可充分吸收两种模式的优点，尝试基于 MOOC 的通识课翻转课堂[38]。数据基础知识、数据获取、数据发布等章节以 MOOC 教学为主，教师的主要精力转为批改作业并定期在线下课堂进行习题讲解和答疑；而数据处理和数据分析部分，则可采取翻转课堂形式，理论与方法部分引导同学们课前学习，课堂则由教师以案例形式带领学生实际操作，同时提供案例视频供学生反复学习和观摩。

5.3　考核方式：个人考核与小组考核的综合

虽然课程采取了综合考核方式来评定学生的最终成绩，但按照学校规定，期末成绩占比不应少于 50%，而本课程的期末考核是以组为单位完成的数据研究报告，虽然也会根据个人在小组中的分工及完成质量综合考量个人的结课成绩，但总体而言难以拉开差距。可能造成最终得分与平时表现不一致的情况；此外，这种考核方式主要考察结课时的数据素养水平，并不能准确反映出其学习效果。因此，今后一方面应该更加细化小组考核，可以建立个人指标和小组指标，前者由组长考核组员[39]，后者则由小组互评和教师考评组成[37]；另一方面亦可结合两次问卷调查结果来综合评估其数据素养能力的提升水平。

5.4　学习效果：学科交叉的混合小组学习模式的构建

虽然课程整体效果良好，但鉴于选课学生的学科差异性，不同学科学生的学习起点不同，认知不同，最终学习效果也有所差异。调查问卷显示，理科生数据素养基础最弱，其次为工学、医学生，人文社科学生基础最好；课程学习之后，各学科学生数据素养水平趋同，理科生跃为最高，14 项指标提升幅度超过 50%，尤其是在数据管理基础知识、数据清洗和分析能力方面。调查问卷体现了学生的自我判断，而小组报告则可反映出其实际应用能力。表 2 表明人文社科学生在数据源及工具利用方面不及理工科丰富，且多学科背景的小组团队普遍得分较高。因此，在国家加快培养交叉学科人才的背景下[40]，充分利用学生学科多样性这一通识课特点，有意识地引导学生在课程早期构建跨学科学习团队，通过学科交融、知识分享来开阔视野，促进学习，在提升数据素养的同时也培养跨学科思维。

参考文献：

[1] 郭华东. 问渠哪得清如许，为有源头活水来——《中国科学数据》发刊词[J]. 中国科学数据，2016(1)：2－5.

[2] POTHIER W G, CONDON P B. Towards data literacy competencies: business students, workforce needs, and the role of the librarian[J/OL]. Journal of business & finance librarianship, 2019 [2020－08－28]. https://www.tandfonline.com/doi/full/10.1080/08963568.2019.1680189.

[3] SHORISH Y. Data information literacy and undergraduates: a critical competency[J]. College & undergraduate libraries, 2015, 22(1): 97－106.

[4] 武汉大学本科生院. 武汉大学关于申报“武大通识3.0”基础通识课、核心通识课和一般通识课的通知[EB/OL]. [2020 – 05 – 15]. http://gec.whu.edu.cn/info/1009/1166.htm.
[5] 孟祥保,常娥,叶兰. 数据素养研究:源起、现状与展望[J]. 中国图书馆学报, 2016, 42(2): 109 – 126.
[6] QIN J, D'IGNAZIO J. Lessons learned from a two-year experience in science data literacy education[C/OL] //International association of scientific and technological university libraries, 31st annual conference 2010 [2020 – 08 – 28]. https://docs.lib.purdue.edu/iatul2010/conf/day2/5/.
[7] CARLSON J, JOHNSTON L, WESTRA B, et al. Developing an approach for data management education: areport from the data information literacy project[J/OL]. The international journal of digital curation, 2013, 8(1)[2020 – 08 – 29]. http://www.ijdc.net/article/view/8.1.204/306.
[8] CARLSON J, JOHNSTON L R. Data information literacy: librarians, data, and the education of a new generation of researchers [M]. West Lafayette: Purdue University Press, 2015.
[9] LISA JOHNSTON J J. Steal this idea: a library instructors' guide to educating students in data management skills[J]. College & research libraries news, 2014, 75(8): 431 – 434.
[10] ZILINSKI L A S N. Thinking critically about data consumption: creating the data credibility checklist[C/OL]//Proceedings of the American Society for Information Science and Technology[2020 – 08 – 28]. https://docs.lib.purdue.edu/lib_fsdocs/89/.
[11] ICPSR. Guide to social science data preparation and archiving [EB/OL]. [2020 – 08 – 20]. https://www.icpsr.umich.edu/web/pages/deposit/guide/index.html.
[12] New England collaborative data management curriculum[EB/OL]. [2020 – 08 – 28]. https://library.umassmed.edu/resources/necdmc/index.
[13] Open access post-graduate teaching materials in managing research data in archaeology[EB/OL]. [2020 – 08 – 28]. https://archaeologydataservice.ac.uk/learning/DataTrain.xhtml.
[14] DATA TREE[EB/OL]. [2020 – 08 – 28]. https://datatree.org.uk/.
[15] 胡卉,吴鸣,陈秀娟. 英美高校图书馆数据素养教育研究[J]. 图书与情报, 2016(1): 62 – 69.
[16] ZILINSKI L, SAPP NELSON M R, VAN EPPS A S. Developing professional skills in STEM students: data information literacy[EB/OL], [2020 – 08 – 20]. http://dx.doi.org/10.5062/F42V2D2Z.
[17] ELIZABETH STEPHENSON P S C. Incorporating data literacy into undergraduate information literacy programs in the social sciences: a pilot project[J]. Reference services review, 2007, 35(4): 525 – 540.
[18] 胡卉,吴鸣,陈秀娟. 加拿大高校图书馆数据素养教育模式[J]. 图书情报工作, 2016, 60(8): 53 – 58.
[19] 禹俊杰. 国外高校图书馆数据素养教育实践与思考[D]. 太原:山西大学, 2018.
[20] 胡卉,吴鸣. 国外图书馆数据素养教育最佳实践研究与启示[J]. 现代情报, 2016, 36(8): 66 – 74.
[21] 郝媛玲,沈婷婷,高珊. 高校数据素养教育实践的思考和建议——基于哈佛大学案例和我国图书情报人员访谈的分析[J]. 图书情报工作, 2015, 59(12): 44 – 51.
[22] 张群,刘玉敏. 高校图书馆科学数据素养教育体系模型构建研究[J]. 大学图书馆学报, 2016, 34(1): 96 – 102.
[23] 张明海,周艳红. 大数据时代大学生数据素养教育的目标定位及体系构建[J]. 图书馆, 2016(10): 84 – 88.
[24] 张群,刘玉敏. 面向研究生的高校图书馆科学数据素养教育研究[J]. 大学图书馆学报, 2017, 35(3): 93 – 98.
[25] 陈明星. 研究生数据素养教育内容框架研究[D]. 南京:东南大学, 2017.
[26] 张斌,刘三女牙,刘智,等. 面向大数据的师范生数据素养课程体系构建研究[J]. 中国远程教育, 2018(4): 62 – 68.
[27] 魏来,王思明. 我国高校图书馆数据素养课程内容构建研究[J]. 情报资料工作, 2018(6): 90 – 95.
[28] 焦皎. 大数据时代高校图书馆数据素养教育研究[J]. 图书馆工作与研究, 2020(3): 81 – 84.
[29] 管进,陈文凯,欧春梅. 基于数据素养教育的文献检索课教学模式探索[J]. 情报探索, 2017(9): 48 – 51.
[30] 张新红. 国内外高校图书馆科学数据素养教育比较分析[J]. 图书馆学研究, 2018(11): 9 – 13.
[31] 宋甲丽,程结晶. 高校图书馆科学数据素养教育现状调查及建议[J]. 图书馆学研究, 2018(19): 2 – 9.
[32] 宋甲丽. 高校图书馆科学数据素养教育研究[D]. 曲阜:曲阜师范大学, 2018.
[33] 武汉大学. 武汉大学2019届毕业生就业质量报告[EB/OL]. [2020 – 05 – 16]. https://info.whu.edu.cn/info/1781/184351.htm.
[34] 张娟,杨文建. 国外数据素养教育研究及其启示[J]. 数字图书馆论坛, 2020(4): 60 – 67.
[35] “共享杯”大学生科技资源共享服务创新大赛[EB/OL]. [2020 – 05 – 31]. http://share.escience.net.cn/nav/index.
[36] “慧源共享.数据悦读”第二届高校开放数据创新研究大赛[EB/OL]. [2020 – 05 – 20]. http://www.huiyuan.sh.edu.cn/competition/#/home.
[37] 宋海艳,余晓蔚,潘卫. 高校信息素养通识教育课程实践探索与思考[J]. 图书馆学研究, 2014(7): 11 – 14.
[38] 温娜,吕海霞,李宝铭. 基于MOOCs的高校通识课翻转课堂教学模式探索[J]. 教育教学论坛, 2019(31): 93 – 95.
[39] 徐菊香. 信息素养通识课翻转课堂的设计与实践[J]. 图书情报知识, 2018(4): 119 – 128.
[40] 王之康. 交叉学科将成为我国第14个学科门类[N]. 中国科学报,2020 – 08 – 11(6).

作者贡献说明:

刘霞:数据分析、论文撰写与修改;
方小利:数据分析、论文修改;
郑怡萍:问卷设计、结果处理。

Construction and Considerations of Data Literacy General Course of Wuhan University for Undergraduates

Liu Xia Fang Xiaoli Zheng Yiping
Wuhan University Library, Wuhan 430072

Abstract: [Purpose/significance] The purpose of this paper is to provide conferences for colleagues of related areas based on the sharing construction process of data literacy general course system of Wuhan University. [Method/process] The teaching objectives, contents, characteristics and affect of the course "data literacy and data utilization" which was carried out under the background of "Wuhan University general education 3.0" were analyzed by case study, statistics analysis and data mining visualization methods, and the rationality of the course was exploited. [Result/conclusion] The teaching methods and practices in and after classes can well stimulate the enthusiasm of the students, and improve the learning effect. Moreover, it can be find from course assessment and the questionnaires before and after the curriculum and the feedback after the course that the students' data literacy has been significantly improved through learning in the course.

Keywords: data literacy general course undergraduates course construction Wuhan University